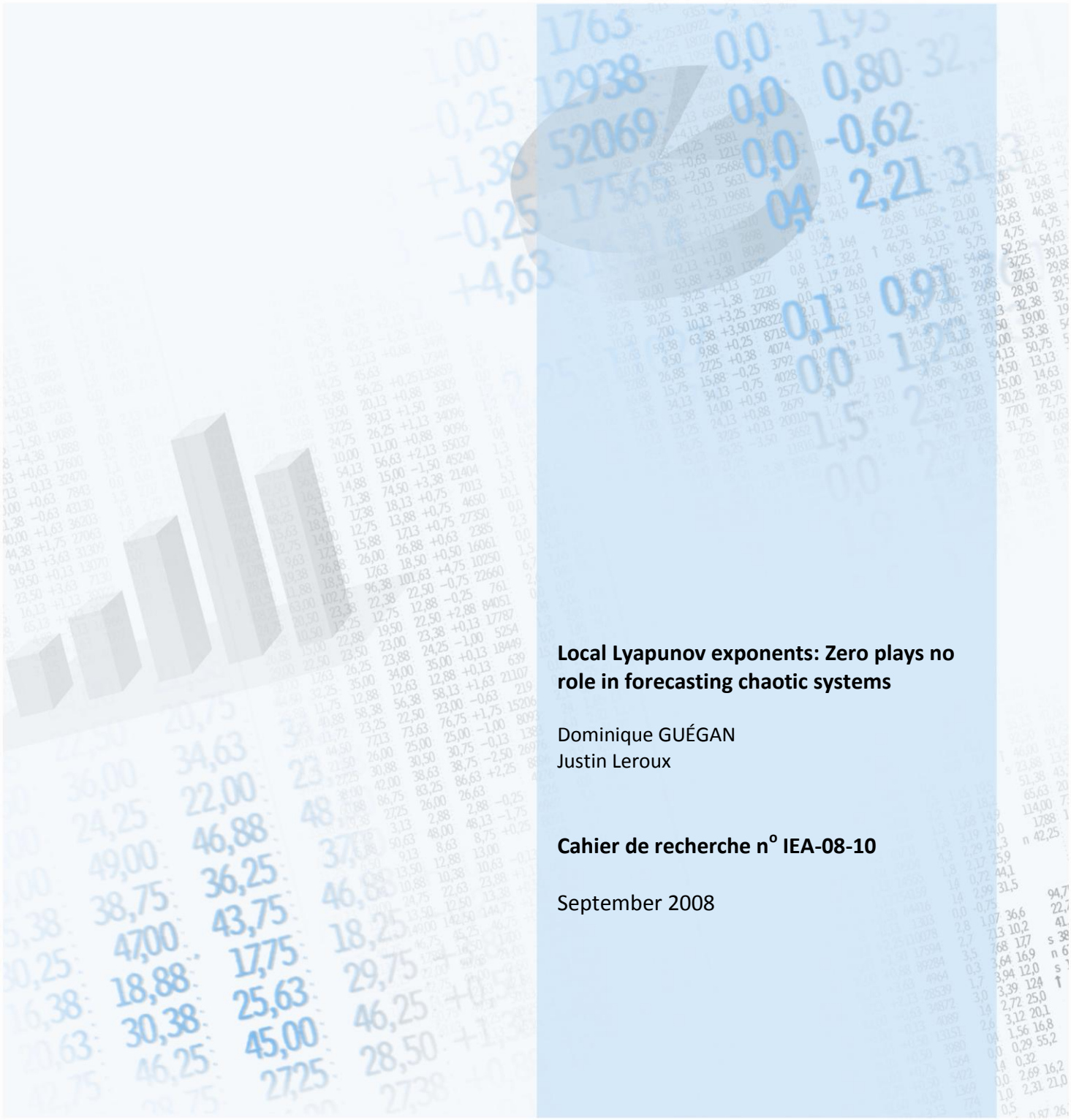
HEC MONTRÉAL
Institut
d'économie appliquée



Local Lyapunov exponents: Zero plays no role in forecasting chaotic systems

Dominique GUÉGAN
Justin Leroux

Cahier de recherche n° IEA-08-10

September 2008

Institut d'économie appliquée
HEC Montréal
3000 chemin de la Côte-Sainte-Catherine
Montréal (Québec) H3T 2A7
Canada
http://www.hec.ca/iea
iea.info@hec.ca

ISSN : 0825-8643

Local Lyapunov exponents: Zero plays no role in forecasting chaotic systems

Dominique Guégan* and Justin Leroux†

September 25, 2008

Abstract

We propose a novel methodology for forecasting chaotic systems which uses information on local Lyapunov exponents (LLEs) to improve upon existing predictors by correcting for their inevitable bias. Using simulated data on the nearest-neighbor predictor, we show that accuracy gains can be substantial and that the candidate selection problem identified in Guégan and Leroux (2009) can be solved irrespective of the value of LLEs. An important corrolary follows: the focal value of zero, which traditionally distinguishes order from chaos, plays no role whatsoever when forecasting deterministic systems.

Keywords: Chaos theory - Lyapunov exponent - Lorenz attractor - Rössler attractor - Monte Carlo Simulations.

JEL: C15 - C22 - C53 - C65.

*Paris School of Economics, CES-Université Paris-I Panthéon Sorbonne, 106-112 boulevard de l'Hôpital, 75013 Paris, France. (Email: Dominique.Guegan@univ-paris1.fr)

†Institute for Applied Economics, HEC Montréal and CIRPÉE, 3000 chemin de la Côte-Ste-Catherine, Montréal, QC H3T 2A7, Canada. (Email: justin.leroux@hec.ca, Fax: (1) 514-340-6469, *corresponding author*)

1 Introduction

When taking a deterministic approach to predicting the future of a system, the main premise is that future states can be fully inferred from the current state. Hence, deterministic systems should in principle be easy to predict. Yet, some systems can be difficult to forecast accurately: such chaotic systems are extremely sensitive to initial conditions, so that a slight deviation from a trajectory in the state space can lead to dramatic changes in future behavior. We propose a novel methodology for forecasting deterministic systems and illustrate how our methodology can be used to improve upon the nearest-neighbor predictor, but the same intuition can be applied to any non-parametric predictor (such as methods based on kernels, radial functions, neural nets, wavelets, etc.; see [1] and [2]) as it corrects for their inevitable bias by incorporating additional information on the local chaoticity of the system via the so-called local Lyapunov exponent (LLE). To the best of our knowledge, while several works exist on the forecasting of chaotic systems (see, e.g., [3], [4], [5], [6], [7], [8]), none exploit the information conveyed by the LLE. The general intuition behind the methodology we propose can be viewed as a complement to existing forecasting methods, and can be extended to chaotic time series. For illustrative purposes, we describe how our methodology can be used to improve upon the well-known nearest-neighbor predictor on two deterministic systems.

The nearest-neighbor predictor has proved to be a simple yet useful tool for forecasting chaotic systems (see [9]). In the case of a one-neighbor predictor, it takes the observation in the past which most resembles today's state and returns that observation's successor as a predictor of tomorrow's state. The rationale behind this nearest-neighbor predictor is quite simple: given that the system is assumed to be deterministic and ergodic, one obtains a sensible prediction of the variable's future by looking back at its evolution from a similar, past situation. For predictions more than one step ahead, the procedure is iterated by successively merging the predicted values with the observed data.

The nearest-neighbor predictor performs reasonably well in the short run but is not satisfactory for even medium-run predictions ([10], [11]). The generally accepted intuition being that the two trajectories (of the current state and of its neighbor) will have separated significantly by then, and the nearest neighbor's medium-run future will have little to do with the future one is trying to predict. Intuitively, this failure to perform well in the medium run arises mainly from the fact that short-run predictions are not accurate enough to withstand the complex dynamics of the system and to remain accurate after being iterated over a period of time of significant length. We argue that this lack of accuracy is inherent to the prediction method itself because the nearest neighbor on which predictions are based can never exactly coincide with today's state (or else the underlying process, being deterministic, would also be periodic and, thus, trivially predicted).

In previous work (Guégan and Leroux, 2009) we propose a novel methodology which aims at correcting the above shortcoming by incorporating information carried by the system's LLE into the prediction. Our methodology yields

two possible candidates, potentially leading to significant improvements over the nearest neighbor predictor, provided one manages to solve the selection problem, which is the issue we address here. We develop a systematic method for solving the candidate selection problem and show, on two known chaotic systems, that it yields statisfactory results (close to a 100% success rate, for the Rössler attractor).

The rest of the paper is organized as follows. In Section 2, we recall the methodology developed in Guégan and Leroux (2008) on the use of LLEs in forecasting and introduce the candidate selection problem. In Section 3, we solve selection problem and show using simulated chaotic systems that the size of the LLEs plays no role in the selection problem. However, the size of the LLEs does matter for the success rate of our selection algorithm and has an impact on the size of errors. These findings are presented in Section 4.

2 Methodology

Consider a one-dimensional series of T observations from a chaotic system, $(x_1, ...x_T)$, whose future values we wish to forecast. Recall that a chaotic system is characterized by the existence of an attractor in a d-dimensional phase space (see [16]), where $d > 1$ is the embedding dimension.[1] A possible embedding method involves building a d-dimensional orbit, (X_t), with $X_t = (x_t, x_{t-\tau}, ..., x_{t-(d-1)\tau})$. For the sake of exposition, we shall assume $\tau = 1$ in the remainder of the paper.

By definition, the local Lyapunov exponent (or LLE) of a dynamical system characterizes the rate of separation of infinitesimally close points of an orbit. Quantitatively, two neighboring points in phase space with initial separation δX_0 are separated, t periods later, by the distance:

$$\delta X \approx \delta X_0 e^{\lambda_0 t},$$

where $|\cdot|$ represents the modulus of the considered vectors and λ_0 is the (largest) LLE of the system in the vicinity of the initial points. Typically, this local rate of divergence (or convergence, if $\lambda_0 < 0$) depends on the orientation of the initial vector δX_0. Thus, strictly speaking, a whole spectrum of local Lyapunov exponents exists, one per dimension of the state space. A dynamic system is considered to be (locally) chaotic if $\lambda_0 > 0$, and (locally) stable if $\lambda_0 < 0$. (see, e.g., [15])

Our goal is to exploit the local information carried by the LLE to improve upon existing methods of reconstruction and prediction. We propose a methodology which builds upon the classical nearest-neighbor predictor, which we now recall. Consider an orbit $(X_1, ..., X_T)$ whose one-step-ahead future, X_{T+1}, we are trying to predict. The nearest-neighbor predictor returns $\hat{X}_{T+1} = X_{i+1}$, where X_i is the element of the orbit with minimal distance to X_T. Because

[1] The choice of the embedding dimension has been the object of much work (see [17] for a survey) and is beyond the scope of this work.

the dynamic system at hand is aperiodic (or else, forecasting would not be an issue), the nearest-neighbor predictor is inevitably biased. Indeed, because $|X_T - X_i| > 0$, it must also be the case that:

$$||\hat{X}_{T+1} - X_{i+1}|| \approx ||X_T - X_i||e^{\lambda_i} > 0, \tag{1}$$

where λ_i can be approximated in practice by the following expression:

$$\hat{\lambda}_i = \ln \frac{||X_{i+1} - X_{j+1}||}{||X_i - X_j||} \qquad \text{with } X_j = \arg\min_{t \neq i,T} ||X_t - X_i|| \tag{2}$$

It follows from Expression (1) that knowing the distance between the predictee and the nearest neighbor as well as the LLE at the nearest neighbor allows us to predict the distance of the predictee's image to the neighbor's image. Note that this is true regardless of the sign of λ_i; i.e., regardless of whether the system is locally chaotic or locally stable.

Moreover, because the orbit considered results from the embedding of a one-dimensional series, we also know all but the first coordinate of $X_{T+1} = (x_{T+1}, x_T, ..., x_{T-d+2})$. Hence, X_{T+1} lies at the intersection of the sphere of radius $||X_T - X_i||e^{\hat{\lambda}_i}$ centered on X_T and the line defined by $\{(z, x_T, ..., x_{T-d+2}) | z \in \mathbb{R}\}$ which, in the Euclidean space, amounts to solving the following polynomial for $z \in \mathbb{R}$:

$$\sqrt{(z - x_{i+1})^2 + (x_T - x_i)^2 + ... + (x_{T-d+2} - x_{i-d+2})^2} - ||X_T - X_i||e^{\hat{\lambda}_i} = 0 \tag{3}$$

Typically, two candidates emerge, $\hat{x}^-_{T+1}$ and $\hat{x}^+_{T+1}$, respectively underestimating and overestimating the true value of observation x_{T+1}(see Figure 1 in the Appendix)[2].

[FIGURE 1 HERE]

3 Solving the selection problem

One difficulty lies in determining when the nearest-neighbor predictor overestimates or underestimates the true value to be predicted. In Leroux and Guégan (2009), we establish that being able to discriminate accurately between $\hat{x}^-_{T+1}$ and $\hat{x}^+_{T+1}$ may significantly improve the accuracy the nearest-neighbor predictor, as we next illustrate. Our goal is to improve on our previous work by developing a systematic selection method to accurately select the best of the two candidates. To do so, we further exploit the information conveyed by the LLE. Indeed, the LLE bieng a measure of local chaoticity of a system (see [13], [14]), it may also yield important clues regarding the regularity of the trajectory

[2] The situation whereby Expression (3) has no real solution would only arise if λ_i had been greatly underestimated, which never occurred to us in practice using Expression (2).

In fact, even "globally chaotic" systems are typically made up of both "chaotic regions" where the LLE is positive and more stable regions where it is negative (Bailey, 1997), as we illustrate in Figures 2 and 3 for the Lorenz and the Rössler system, respectively. Hence, we may expect very stable trajectories where the LLE is small, wheras regions where the LLE is large yield highly unstable behavior.

[FIGURE 2]

[FIGURE 3]

Thus, we have investigated conditioning our selection process on the value of the LLE. Formally, our algorithm can be defined as follows:

$$\begin{cases} \text{If } \lambda_T \leq \bar{\lambda}, \text{ select the "most collinear" candidate} \\ \text{otherwise, select the "least collinear" candidate,} \end{cases} \tag{4}$$

where $\bar{\lambda}$ is an exogenously given threshold value. We abuse terminology slightly and denote by "most collinear" (resp. "least collinear") the candidate $\hat{x}_T = \hat{x}^-_{T+1}, \hat{x}^+_{T+1}$ such that the scalar product $(\hat{X}_{T+1} - X_T) \cdot (X_{i+1} - X_t)$ is minimum (resp. maximum). In words, this algorithm assumes that when the value of the LLE is low, the orbit is relatively smooth, suggesting that the trajectory to be predicted behaves similarly as the nearest neighbor's trajectory. Alternatively, when the LLE is "large", the trajectory is considered to behave erratically, so that the trajectory to be predicted is assumed to differ from that of its nearest neighbor.

Computations of average errors across all values of $\bar{\lambda}$ in the range of the system's LLE actually yield that the best value when $\bar{\lambda}$ is the upper bound of the range (Figure 4). In other words, one is better off alway selecting the most collinear candidate and not conditioning the selection process on the LLE as in Expression (4).

[FIGURE 4]

In the remainder of the paper, we evaluate the performance of the predictor which systematically selects the most collinear candidate. We find (see Table 1) that our procedure almost always selects the best of the two candidate and that improvements upon the NNP, which is well-known for performing very well in the very short run, are substantial. Surprisingly, of all 100 predicitons, the only selection mistake occurs for a negative value of the LLE (-0.1776). Moreover, when making 1000 predictions, our procedure still selects very accurately even-though the range of the LLE is large. These findings confirm the fact that our selection procedure is hardly affected by the sign of the LLE.

Table 1: Rössler system		
Number of predictions	100	1000
Success rate	99%	97.3%
Average error	0.0172	0.0157
NNP average error	0.0279	0.0236
Best candidate average error	0.0172	0.0155
otherSE	0.0590	0.0452
mean LLE (min;max)	0.1979 (-0.5254;0.6573)	0.1302 (-1.2453;0.9198)
mean LLE on mistakes (min;max)	-0.1776 (–;–)	0.2582(-0.4824;0.9198)

With the Lorenz system (Table 2), we find that our procedure selects better over 1000 predictions than it does over 100 predictions. We suspect this is the case because the 100 predictions were made over a more chaotic segment of the orbit, on average. The fact that the average LLE is larger on this segment of the trajectory (0.2756 versus 0.1948) corroborates our suspicion. Hence, the value of the LLE does seem to play a small role in the selection problem. Nevertheless, recall that our results show that conditioning selection on the value of the LLE as in Expression (4) would not lead to an improved success rate.

Table 2: Lorenz system		
Number of predictions	100	1000
Success rate	87%	94.30%
Average error	0.0738	0.0390
NNP average error	0.1075	0.0548
Best candidate average error	0.0689	0.0372
otherSE	0.2041	0.1036
mean LLE (min;max)	0.2756(-0.9861;1.3639)	0.1940 (-1.4353;1.4580)
mean LLE on mistakes (min;max)	0.4685(-0.4020;1.3639)	0.4354 (-0.5142;1.3639)

4 Discussions

In this section, we detail the role of the LLE on the size of errors and on the performance of the selection procedure.

4.1 Role of the LLE on error size

Average of errors:

Table 3: Average size of prediction errors. 1000 predictions.						
	Rössler			Lorenz		
LLE	success	failure	raw 1-NN	success	failure	raw 1-NN
-1.5,1.3	-	-	-	0.0077	-	0.0077
-1.3,-1.1	$6.2538e^{-6}$	-	$6.2538e^{-6}$	-	-	-
-1.1,-0.9	-	-	-	0.0335	-	0.0335
-0.9,-0.7	$6.7189e^{-4}$	-	$7.2159e^{-4}$	0.0286	-	0.0286
-0.7,-0.5	0.0188	-	0.0191	0.0451	0.0169	0.0479
-0.5,-0.3	0.0153	0.0045	0.0161	0.0541	0.005	0.0594
-0.3,-0.1	0.0199	0.0068	0.0224	0.0505	0.0899	0.0626
-0.1,0	0.0428	0.0012	0.0850	0.0477	0.0347	0.0580
0,0.1	0.0106	$5.9019e^{-7}$	0.0168	0.0249	0.027391	0.0365
0.1,0.3	0.0100	0.0214	0.172	0.0342	0.0876	0.0539
0.3,0.5	0.0162	0.0623	0.0219	0.0290	0.0594	0.0626
0.5,0.7	0.0108	0.0170	0.0253	0.0325	0.0847	0.0560
0.7,0.9	-	-	-	0.0236	0.0562	0.0267
0.9,1.1	-	$3.0565e^{-5}$	$5.2792e^{-6}$	0.0152	0.396	0.0273
1.1,1.3	-	-	-	0.0433	0.0632	0.0333
1.3,1.5	-	-	-	0.1594	0.0432	0.0339
Mean	0.0154	0.0266	0.0236	0.0373	0.0667	0.0548

Notice that the size of errors is relatively stable over the range of LLEs when selection is successful. This indicates that our method accurately corrects for the dispersion of neighboring trajectories as measured by the value of the LLE. If this were not the case, one would expect the size of errors to be larger for larger values of LLEs. In fact, errors become large only for values of the LLE near the upper end of their range (above 0.9 for the Rössler attractor, and above 1.1 for the Lorenz attractor). A possible reason for this sudden increase may be that our estimator for the value of the LLEs is not sufficiently robust in regions of high chaoticity. We expect that a more sophisticated estimation method for the LLEs may solve this issue.

Notice that for the Rössler attractor, for values of the LLE in the [-0.5, 0.1] range, the size of errors when failing to select is on average less than when selecting accurately. This apparently surprising observation is actually encouraging, as it indicates that selection mistakes occur mostly when there is little need for correction. Such situations may arise because X_T's nearest neighbor is very close to X_T or, alternatively, when both candidates, $\hat{x}^-_{T+1}$ and $\hat{x}^+_{T+1}$ are very close to x_{i+1} due to space orientation considerations.

4.2 Role of the LLE on success rate

Success rates:

Table 4: Success rates. 1000 predictions.						
	Rössler			Lorenz		
LLE	success	failure	success rate	success	failure	success rate
-1.5,1.3	-	-	-	1	0	100%
-1.3,-1.1	1	0	100%	-	-	-
-1.1,-0.9	-	-	-	3	0	100%
-0.9,-0.7	5	0	100%	3	0	100%
-0.7,-0.5	68	0	100%	67	1	98.51%
-0.5,-0.3	106	2	98.14%	92	1	98.91%
-0.3,-0.1	105	3	97.22%	98	2	97.96%
-0.1,0	46	2	95.83%	47	4	91.49%
0,0.1	79	1	98.75%	61	4	93.44%
0.1,0.3	149	4	97.39%	109	7	93.58%
0.3,0.5	222	8	96.52%	195	8	95.90%
0.5,0.7	192	6	96.97%	223	22	90.13%
0.7,0.9	-	-	-	18	2	88.89%
0.9,1.1	0	1	0%	13	3	76.92%
1.1,1.3	-	-	-	9	2	77.88%
1.3,1.5	-	-	-	4	1	75%
Total	973	27	97.3%	943	57	94.3%

According to Table 4, our selection method performs almost perfectly on the Rössler system. Our method selects perfactly for very low values of the LLE. Selection mistakes start to appear past a value of LLE of -0.5, but the success rate does not seem to fall significantly as the value of the LLE increases up to 0.7. These results are evidence against the common intuition according to which trajectories are more stable, or smoother, where the value of the LLE is small and more irregular for large values of the LLE. For the Lorenz system, however, the success rate falls as the LLE grows large, which is more in line with the common intuition. Interestingly, mistakes occur even for negative values of the LLE, where the system is stable, by definition. Hence, our results suggest that the focal value of $\lambda = 0$, traditionally separating order from chaos, is irrelevant in terms of forecasting.

5 Concluding comments

This follow-up on our encouraging preliminary study ([12]) confirms that the potential gains identified there are achievable. Indeed, the candidate selection problem is no longer an issue thanks to a selection method which, surprisingly, does *not* condition on the value of the LLE. Moreover, thedistribution of pre-

diction errors and success rates we obtain sugegst that the focal value of zero for the LLE, which traditionally separates stability from chaotic behavior, plays no role when it comes to forecasting

The next steps include enhancing predictions via better estimations of the LLE, either by using more neighbors, or thanks to neural network methods (Gençay, 199x). Naturally, our ultimate goal is to evaluate how our method holds up when confronted to real data, and particularly intra-day financial and economic time series.

References

[1] Shintani M and Linton O. Nonparametric neural network estimation of Lyapunov exponents and a direct test for chaos, 2004; 120: 1-33.

[2] Guégan D and Mercier L. Stochastic and chaotic dynamics in high-frequency financial data. Signal Processing and Prediction, Prochazka Eds, 1998; 365-372.

[3] Doerner R, Hübinger B, Martienssen W, Grossman S and Thomae S, Predictability Portraits for Chaotic Motions, Chaos, Solitons and Fractals, 1991;1:553-571.

[4] Rinaldi S, Casagrandi R and Gragnani A, Reduced order models for the prediction of the time of occurrence of extreme episodes, Chaos, Solitons and Fractals, 2001;12: 313-320.

[5] Yousefi S, Weinreich I and Reinarz D, Wavelet-based prediction of oil prices, Chaos, Solitons and Fractals, 2005; 25: 265-275.

[6] Yang HY, Ye H, Wang G, Khan J and Hu T, Fuzzy neural very-short-term load forecasting based on chaotic dynamics reconstruction, Chaos, Solitons and Fractals, 2006;29: 462-469.

[7] Chian ACL, Rempel, EL and Rogers C, Complex economic dynamics: Chaotic saddle, crisis and intermittency, Chaos, Solitons and Fractals, 2006; 29: 1194-1128.

[8] Guerra FA and Coelho LS, Multi-step ahead nonlinear identification of Lorenz's chaotic system using radial basis neural network with learning by clustering and particle swarm optimization, Chaos, Solitons and Fractals, 2009; 35:967-979.

[9] Farmer JD and Sidorowich JJ. Predicting chaotic time series. Physical Review Letters, 1987; 59: 845 – 848.

[10] Ziehmann C, Smith LA and Kurths J. Localized Lyapunov exponents and the prediction of predictability. Physics Letters A, 2000; 271: 237-251.

[11] Guégan D. Les Chaos en Finance: Approche Statistique. Economica Eds.; 2003, 420 pp.

[12] Guégan D and Leroux J. Forecasting chaotic systems: The role of local Lyapunov exponents. Chaos, Solitons and Fractals, 2009, forthcoming.

[13] Abarbanel HDI. Local and global Lyapunov exponents on a strange attractor. Nonlinear Modeling and Forecasting, SFI Studies in the Science of Complexity, Casdagli M and Eubank S Eds, Addison-Wesley, 1992; Proc. Vol. XII: 229-247.

[14] Wolff RCL. Local Lyapunov exponents: looking closely at chaos. Journal of the Royal Statistical Society B, 1992; 54: 353 – 371.

[15] Bailey B. Local Lyapunov exponents: predictability depends on where you are. Nonlinear Dynamics and Economics, Kirman et al. Eds, 1997.

[16] Eckmann JP and Ruelle D. Ergodic theory of chaos and strange attractors. Review of Modern Physics, 1985; 57: 615-656.

[17] Takens F. Estimation of dimension and order of time series. Progress in Nonlinear Differential Equations and their Applications, 1996; 19: 405-422.

[18] Takens F. Detecting strange attractors in turbulence. Lecture Notes in Mathematics, 1981; 898: 366-381.

[19] Lorenz EN. Deterministic non-periodic flow. Journal of Atmospheric Science, 1963; 20: 130-141.

[20] May RM. Simple mathematical models with very complicated dynamics. Nature, 1976; 261: 459-467.

6 Appendix

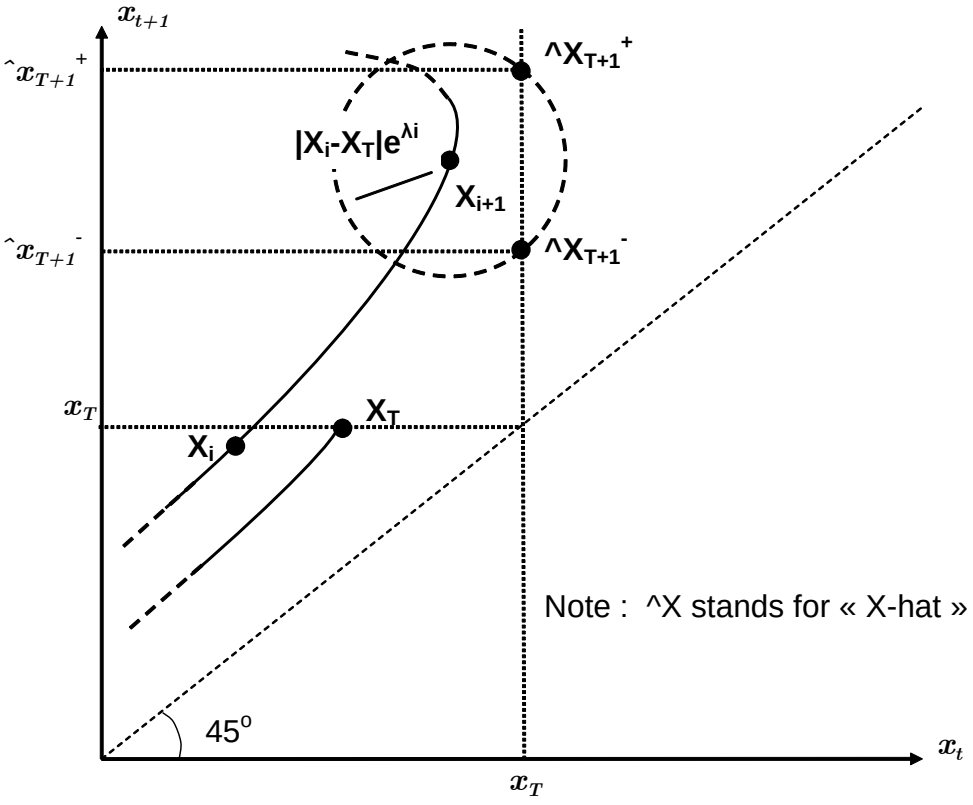


Figure 1: Determination of the two candidates: $\hat{x}_{T+1}^{-}$ and $\hat{x}_{T+1}^{+}$.

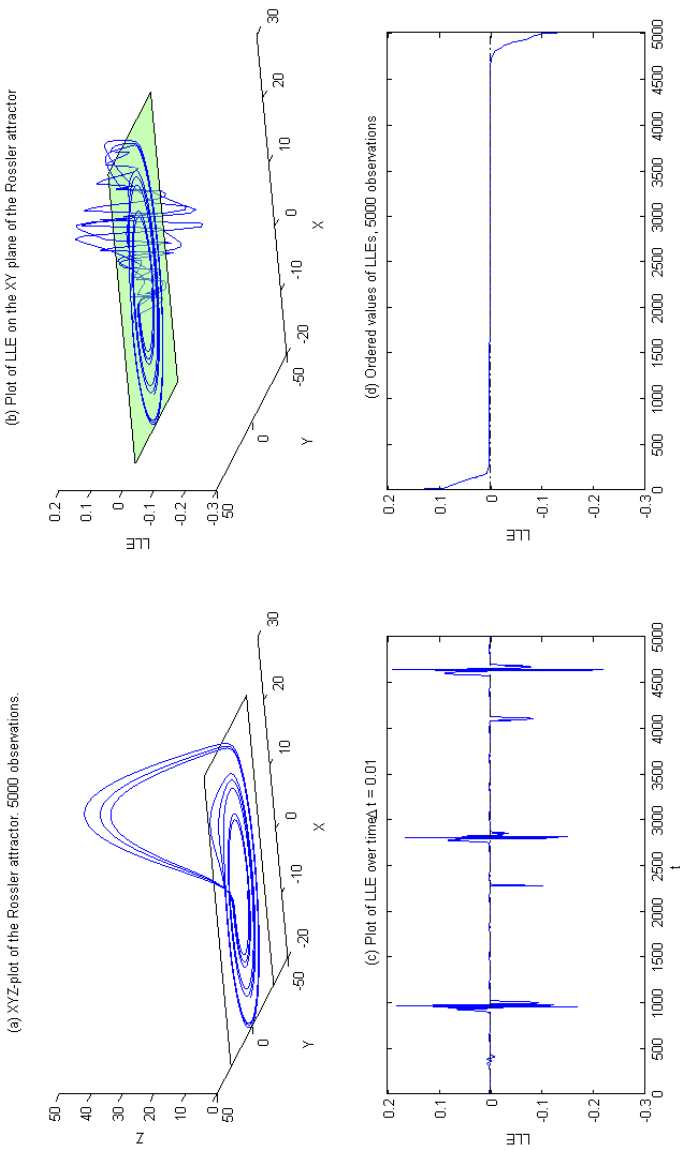


Figure 2: Evolution of LLE for the Rössler system

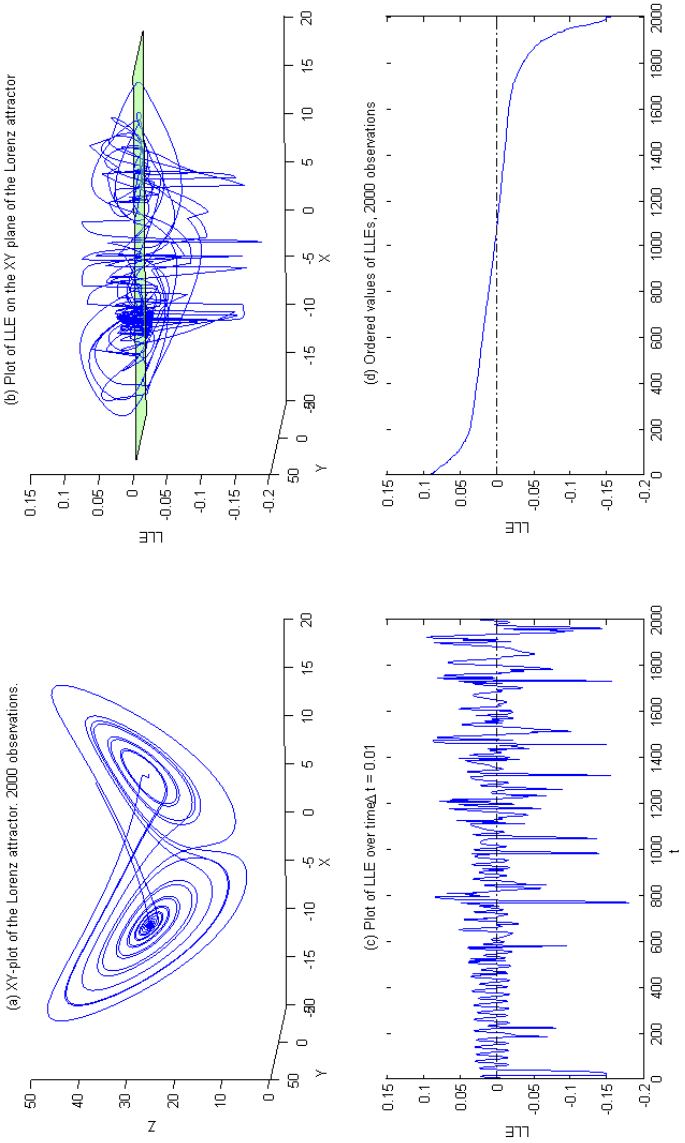


Figure 3: Evolution of LLE for the Lorenz system

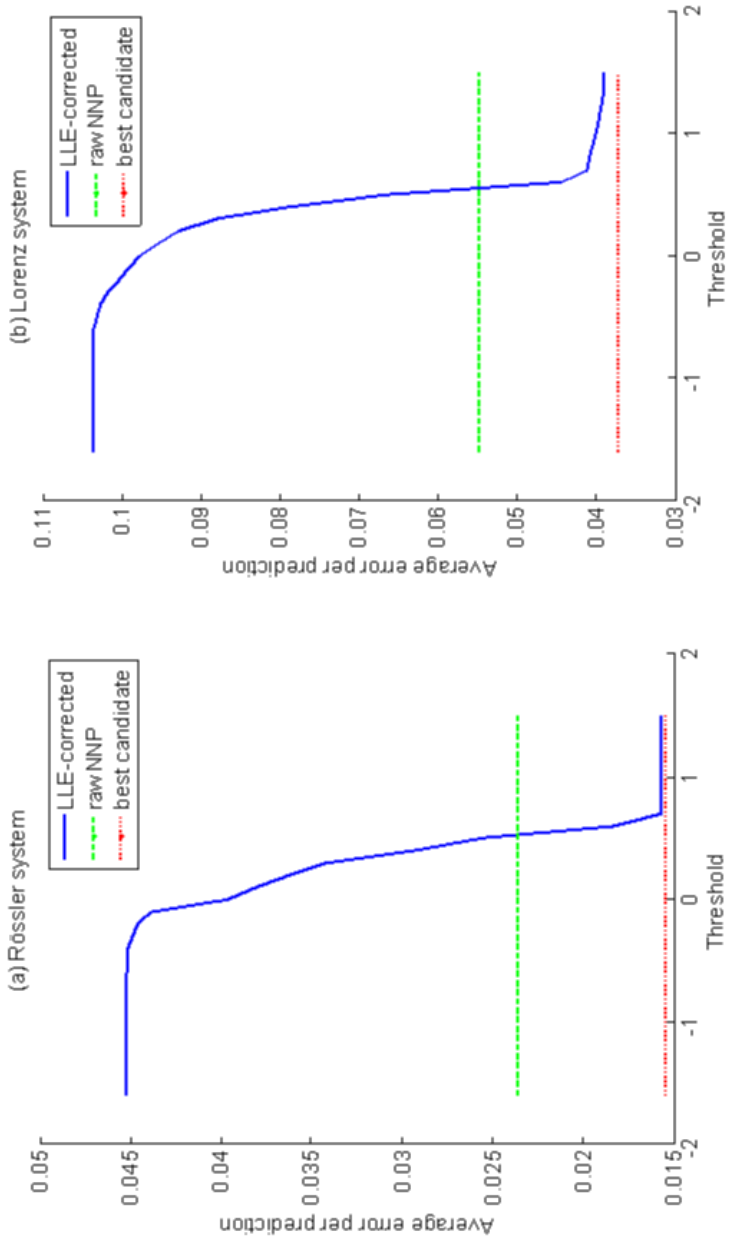


Figure 4: Average error per prediction as a function of threshold $\bar{\lambda}$. 1000 predictions